Kick: Investigating the Use of Kick Gestures for Mobile Interactions

Teng Han[1], Jason Alexander[1], Abhijit Karnik[1], Pourang Irani[2] and Sriram Subramanian[1]

[1]University of Bristol
Bristol, United Kingdom
{han, jason, karnik, sriram}@cs.bris.ac.uk

[2]University of Manitoba
Winnipeg, MB, Canada
irani@cs.manitoba.ca

ABSTRACT
In this paper we describe the use of kick gestures for interaction with mobile devices. Kicking is a well-studied leg action that can be harnessed in mobile contexts where the hands are busy or too dirty to interact with the phone. In this paper we examine the design space of kicking as an interaction technique through two user studies. The first study investigated how well users were able to control the direction of their kicks. Users were able to aim their kicks best when the movement range is divided into segments of at least 24°. In the second study we looked at the velocity of a kick. We found that the users are able to kick with at least two varying velocities. However, they also often undershoot the target velocity. Finally, we propose some specific applications in which kicks can prove beneficial.

Author Keywords
Foot interaction, kicking, Mobile HCI.

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): Interaction styles.

General Terms
Human Factors.

INTRODUCTION
Mobile phones are often used in contexts where users' hands are either too dirty to touch the screen or are covered due to environmental conditions (weather, sterile environments, etc.). In such contexts user's can at best hold their phone but often cannot use touch for even basic interactions. One interesting approach to touchless interaction is to use foot gestures (e.g. [3, 5, 7]). For example, foot gestures are available/doable in cold weather, where users cannot remove their gloves to operate a phone, or on farms where farmers with dirty hands wish to use their phones to research fertilizers while in their fields.

Foot movement is a robust input method for many tasks (i.e. driving, dancing, running) leading researchers to study foot gestures in various contexts [3]. Foot input has been used for selecting menu items [7], for game interaction [5] and to control a device [3]. Multitoe [2] enabled identity tracking from foot imprints and with computer vision techniques users can play a mock football game on a PDA using their own foot [5].

More closely related to our work is that of applying foot gestures, such as ankle rotations for discrete selection [3, 7]. This prior work mapped specific foot movements to an interaction. We instead investigate interactions that take advantage of the multiple degrees-of-freedom available in a foot gesture, such as a kick (Figure 1). This also has the benefit of being easily learned and adopted by users for a variety of tasks, since foot gestures are most likely to be used occasionally when the hands are busy.

Figure 1. Kicking as a method of interaction with mobile applications: (left) Directional 'kick' gesture, (right) Velocity based 'kick' gesture.

This paper investigates the dexterity of using '*kicking*' as a foot gesture in mobile interaction through two studies. The first reveals that kicks are precise enough to distinguish up to five different directions in front of the user. The second shows that users undershoot their goal targets in velocity-based kicking movement. The main contributions of this paper are: a) an exploration of kicking as an interaction technique; and b) an investigation of the dexterity of kick direction and velocity for use in interactive tasks.

KICKING
We first explain our use of the term 'kick' and then describe the details of our kick detection algorithm.

Kick Gesture
The meaning of a 'kick' is usually placed within a context, such as 'to kick a ball', 'a round house kick', or 'kicking off

ones shoes'. Such actions require a large number of muscle groups. While kicking takes on different forms we limit our definition to those actions that seem socially acceptable and practical to perform on a mobile device. As such, a 'kick' gesture in our context consists of moving one's foot forward, left or right, and thus only using the muscles of the lower leg (tibialis anterior or shin muscle, and calve muscles, along with foot and ankle tendons).

Detecting a Kick

When a user kicks an item (e.g. a soccer ball), up to six muscle groups in the leg cooperate to perform the action in a controlled manner [1], with up to three muscles for a short range kick. As a result, kicking can be extremely expressive. However, discretizing a kick into its component actions for interaction would limit the amount of expressivity it can afford. Instead other alternatives are needed for interpreting a kick to fully harness its power.

Unlike prior work that primarily discretized foot movement for extracting taps or ankle rotations, we instead interpret kick gestures by capturing a complete foot gesture with a camera and feeding it into a physics engine. The advantage of this approach is that it allows the users to formulate their own impression of what to do with a kick. This also permits a wider range of mappings that would not be possible by simply breaking down a complex movement into subparts.

To detect a kick we used an Xbox Kinect camera (placed ~3.5m in front of the users) and computer vision algorithms to extract the foot gesture information. The depth camera extends the operating dimensions from 2D to 3D space so that users can freely perform any foot gestures. This solution was selected as while such depth cameras are not currently available on mobile devices, it is likely they will be in the near future. Furthermore, we need a resilient detection system to perform our evaluations and to demonstrate our concept. The Open Natural Interaction (OpenNI) library provides a user skeleton model and helps to track users' bodies easily. We extract the users' right foot coordinate information, and detect kick gestures from it using the following algorithm.

The system runs at ~25 fps and even when the user is still, it detects minor movement (change in the spatial coordinates of the foot). Thus we characterize the 'kick' gesture as significant movement of the foot (above a preset threshold). Natural movements like shuffling or adjusting balance from one foot to another are gradual and not detected as a 'kick' by our filter criteria. A normal 'kick' gesture was found to take more than 0.35s to execute during preliminary studies. So we also filtered out any actions which did not last longer than nine frames (~0.35s).

The spatial position of the foot was recorded in every frame during a kick. To detect the direction (with respect to the ground plane), we ignored the vertical axis data. Using the method of least squares the system calculates a line of best fit with the foot trajectory (again with respect to the ground plane). The slope of this equation provides the direction. The vertical axis data along with the direction is used to detect the 'kick' velocity. Velocity is converted to force (for the physics engine) using a linear equation.

USER STUDY 1: A KICK IN THE RIGHT DIRECTION

To establish a basic understanding of kick interactions and to determine the accuracy with which users can control their kicks, we conducted an initial user study. This study had two goals: 1) to establish the feasibility of using kick gestures to interact with a mobile device and 2) to determine in how many distinct directions users can accurately perform a kicking gesture.

To do this, we designed a football game (see Figure 2a) that required users to kick the virtual ball into a highlighted target goal. We began with a football representation so that participants could easily associate their physical actions with a familiar (virtual) representation. Moving immediately to an abstract representation such as a menu could potentially increase the learning time and break the inherent metaphor transferred by the kick action.

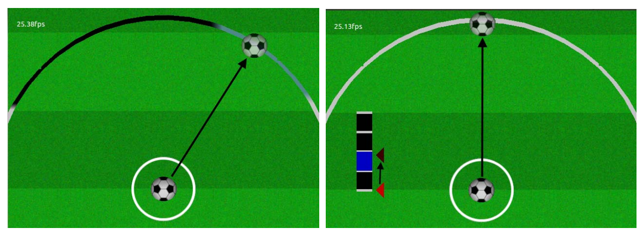

Figure 2. User studies: (left) Using direction of 'kick' to propel football towards the (blinking) grey goal (right) Using velocity of 'kick' to propel football while keeping it within the blue-band target range.

Participants and Apparatus

Eight volunteers (three female), between the ages of 20 and 35 years, participated in the experiment. All participants were right leg dominant. Participants held a 7" display at approximately waist height—they looked down onto the screen as if they were looking down to the ground to kick a football (as illustrated in Figure 1a). Their gestures were tracked by a Microsoft Xbox Kinect, as described earlier.

Experimental Interface and Tasks

The experimental interface (Figure 2a) filled the 7" display, which had a resolution of 800×480 px. The ball always began in the same starting location. A black semi-circle indicated the 120° range where the ball may be directed, while the target goal was the blinking color strip on this semi-circle. To perform a trial, the user kicked in the direction of the goal. The ball would then move in the direction of the kick until it had passed over the semi-circle. In this study, the velocity of the kick was ignored and so the ball always moved at a constant velocity.

Design and Procedure

The experiment consisted of a practice session (five kicks) and the recorded trials. Participants were free to rest at any point during the experiment. The single independent variable was the *goal width*, measured in degrees. This took values of 40°, 30°, 24°, 20°, 17.1°, 15°, 13.3° and 12° (derived from having 3–10 divisions around the 120° semi-circle). We recorded the position at which the ball crossed the semi-circle, giving us a success/miss record and an angular measure of error (when a miss occurred). Erroneous kicks were not repeated.

Post-practice, all participants moved from the largest goal width through to the smallest goal width. Participants were required to kick the ball into each possible goal position three times (the positions were provided in a random order). So for example when the number of divisions were 5 (angular width 24°) the total number of trials per participant was 5×3 = 15. The total number of trials per participant was 156 (9+12+15+18+21+24+27+30).

Results and Discussion

The total number of misses was 438 out of 1248 trials. There was a significant effect of goal width on accuracy ($p < 0.01$, $F_{(7,49)} = 24.3$). 40° divisions had an accuracy of 96%, while 12° divisions had 46%. Five divisions, with an angular width of 24°, had an accuracy of 88%, a value we would recommend for use in further kick-based selections.

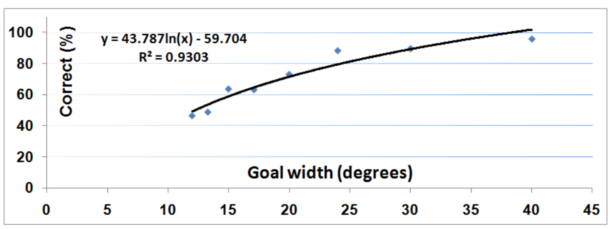

Figure 3. Study 1 results – accuracy of directional kicks

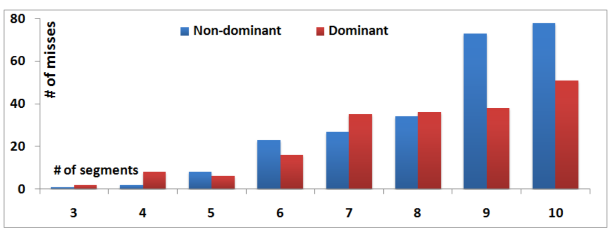

Figure 4. Number of misses in the dominant and non-dominant sides for each number of divisions.

We also found that there was a total of 192 misses on the dominant side of the target goal while there were 246 misses on the non-dominant side (where the dominant side of the target goal is the halve of the goal that is on the same side as the user's dominant leg). Figure 4 shows a breakdown of these misses for each number of divisions, suggesting that there are fewer misses when participants tried to kick the ball towards their dominant side.

This study has shown that participants can accurately direct their kicks into divisions spaced 24° apart. Users are also more accurate on their dominant side than the non-dominant side. We now wish to investigate whether it is feasible to use kick velocity as an interaction parameter.

USER STUDY 2: A KICK WITH THE RIGHT VELOCITY

Having gained an understanding of users' abilities to correctly direct a kick, we wished to explore how users can control kick velocity. This will allow us to judge the feasibility of using velocity as an input dimension. This study was run using the same setup as User Study 1.

Participants and Apparatus

Eight volunteers (two female) between the 20 35 years old participated in the experiment, seven of whom were right leg dominant. We used the same hardware setup as Study 1.

Experimental Interface and Tasks

A football game was again employed, this time with users having to kick the virtual ball with a specified velocity. We disregarded the direction of the kick, but encouraged participants to kick straight instead of sideways. The velocity of each kick was visualized on-screen by a velocity meter, as shown in Figure 2b. The velocity meter was updated in real-time as the participants performed a kick, providing them feedback about the kick velocity. In each task the user was shown a target range of velocities. The participant's goal was to then kick the ball such that the kick velocity would be within the target range.

Design and Procedure

The experiment consisted of a practice session (6 kicks), a user calibration session (users' performed 6 kicks to register maximum and minimum velocities: 3 for minimum velocity (V_{MIN}) and 3 for maximum velocity (V_{MAX})) and the recorded trials, with participants free to rest at any point. The single independent variable was the *number of velocity divisions*. The velocity divisions were spread equally over the calibration range (between V_{MAX} and V_{MIN}) for the user. The number of divisions varied between 2 to 4. For each trial we recorded the velocity of the kick, giving us a success/failure record and a measure of error when a kick was too fast or too slow. Erroneous kicks were not repeated. All participants moved from the smallest number of divisions (2) to the largest (4). For each value of number of velocity divisions, participants were required to kick the ball within the velocity values of the required segment three times. For example, in the four divisions' case, the participants completed 4×3=12 tasks while they completed 6 and 9 tasks for the two and three divisions respectively. The target division was provided in a random order.

Results and Discussion

There was a significant difference between the accuracy of the user and the number of kick divisions ($p < 0.05$, $F_{(2,14)} = 7.8$ when analyzed using a Univariate ANOVA with participants as random factors).

The two divisions condition was the best (statistically significant in post-hoc pair-wise comparisons with $p < 0.01$) with an overall accuracy of 87.5% (42 out of 48 trials). We found no statistical difference between 3 and 4 divisions. Figure 5 shows the average percentage accuracy for each division. We further analyzed the data to see if users were undershooting or overshooting the target. Our results showed that users were often undershooting the target. Of the 73 cases where they failed to reach the desired level, 49 cases were undershot while only 24 were overshoot errors. Figure 6 shows the breakdown of errors over each division.

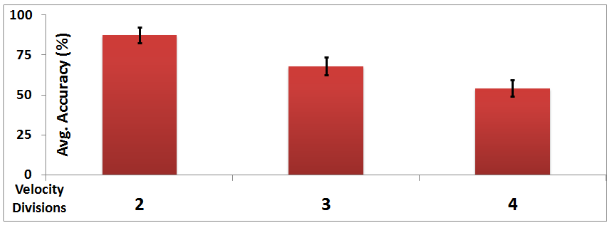

Figure 5: Average accuracy of tasks (expressed in %) for each division. We use average instead of actual numbers as the number of trials for each division was different.

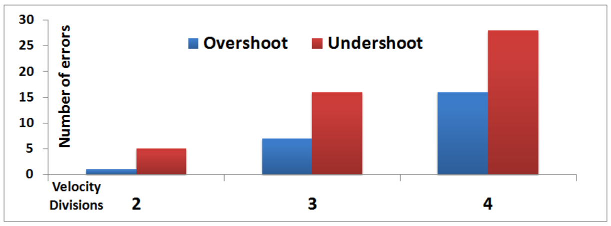

Figure 6: Total undershoots and overshoots per division.

DISCUSSION

Based on the results of our studies, we consider several possibilities where the kick gesture could be useful.

Kick to flick: Our studies show that it is harder to precisely control the velocity of the kick gesture. However, the users can remember two broad ranges of velocity. Thus the flick action is easily interpreted with the kick. The flick action on a menu would be triggered by a higher velocity kick. The slower velocity kick and direction of kick can be used to provide additional control as required for a Superflick [6].

Kick to navigate: A circular contextual marking-menu like the hierarchical marking menu [8] is a good example of a menu that can work well with the kick action. A naïve user would receive feedback for slower directional gestures as they progress through sub menus. However an expert user, who remembers the menu layout, could navigate quickly in a single continuous gesture. Our studies suggest that five divisions at each level of the menu would be helpful for retaining the selection accuracy. The velocity of the kick can be leveraged for navigation in a hierarchical step.

Kick to zoom: Igarashi et al.[4] demonstrated speed-dependent zooming actions. The kicking is well suited for to this interaction as the user has two kick velocities that can be used to control the interface's zoom action.

Improving kick distance: Study 2 showed that users tend to undershoot the target. Pressure based interaction techniques have successfully applied quadratic mapping functions to increase the number of selection levels. Future studies could look at such non-linear mapping functions between the kick gesture speed and document movement speed to further increase the step size for direction.

Our study only explored stationary interactions and future research could look at kick gestures integrated with a walking action. In such cases relative alignment of the phone to the body can change. Further areas for consideration include the social acceptability of such gestures and the minimum space requirements for performing these gestures.

CONCLUSION

In this paper, we investigate the dexterity of using kicking as an effective foot gesture in mobile interaction contexts. The two studies indicate that the users can distinguish between and control up to five directions and two kicking velocities. A mobile interface using kick gestures as input is feasible if due care is taken not to make it complex. Simple interactions like choosing menus, scroll lists and navigating maps can be easily adapted for the kick gesture.

ACKNOWLEDGEMENTS

This work was funded jointly by EPSRC (grant number EP/G058334/1) and MobileVCE (www.mobilevce.com) as part of the User Interactions for Breakthrough Services research program.

REFERENCES

1. Man-Systems Integration Standards, NASA-STD-3000, Revision B, (1995)
2. Augsten, T., Kaefer, K., Meusel, R., Fetzer, C., Kanitz, D., Stoff, T., Becker, T., Holz, C. and Baudisch, P. Multitoe: high-precision interaction with back-projected floors based on high-resolution multi-touch input. UIST '10. New York, NY, USA, ACM: 209-218, (2010)
3. Crossan, A., Brewster, S. and Ng, A. Foot Tapping for Mobile Interaction. BCS HCI '10, (2010)
4. Igarashi, T. and Hinckley, K. Speed-dependent automatic zooming for browsing large documents. UIST '00. San Diego, CA, USA, ACM: 139-148, (2000)
5. Paelke, V., Reimann, C. and Stichling, D. Foot-based mobile interaction with games. ACE '04. Singapore, ACM: 321-324, (2004)
6. Reetz, A., Gutwin, C., Stach, T., Nacenta, M. and Subramanian, S. Superflick: a natural and efficient technique for long-distance object placement on digital tables. GI '06. Quebec, Canada, Canadian Information Processing Society: 163-170, (2006)
7. Scott, J., Dearman, D., Yatani, K. and Truong, K. N. Sensing foot gestures from the pocket. UIST '10. New York, NY, USA, ACM: 199-208, (2010)
8. Zhao, S. and Balakrishnan, R. Simple vs. compound mark hierarchical marking menus. UIST '04. Santa Fe, NM, USA, ACM: 33-42, (2004)